BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 3: e37257
doi: 10.3897/biss.3.37257

OPEN ACCESS

Conference Abstract

Making LTER Data FAIR: A workbench using DEIMS datasets and GBIF Tools

Francisco Pando[‡], Francisco Bonet[§]

‡ Real Jardin Botanico -CSIC, Madrid, Spain
§ Universidad de Córdoba, Córdoba, Spain

Corresponding author: Francisco Pando (pando@rjb.csic.es)

Received: 13 Jun 2019 | Published: 19 Jun 2019

Citation: Pando F, Bonet F (2019) Making LTER Data FAIR: A workbench using DEIMS datasets and GBIF Tools. Biodiversity Information Science and Standards 3: e37257. https://doi.org/10.3897/biss.3.37257

Abstract

DEIMS-SDR (Dynamic Ecological Information Management System - Site and dataset registry, Wohner et al. 2019) is one of the largest repositories of long-term ecological research (LTER) datasets. It provides sophisticated searching tools by metadata elements and identifiers for all the 930 contained datasets, most of them from European sites. Whereas datasets' metadata are highly structured and searchable, datasets themselves have little standardization in terms of content, identifiers or license, making data integration difficult or cumbersome. Adopting the data FAIR guiding principles(Wilkinson et al. 2016) for LTER data would result in better data integration and reutilization to support knowledge discovery and innovation in ecological research.

The Global Biodiversity Information Facility (GBIF 2019a). is the largest repository of species distribution data in the world, providing access to more than a billion records from over 43,000 datasets. GBIF is a good example of FAIR principles implementation: GBIF data is highly standardized, using Darwin Core (Wieczorek et al. 2012) for data and ecological metadata language (EML, Fegraus et al. 2005) for metadata, allowing record-level search; and has implemented globally unique and persistent identifiers for datasets and downloads. Relevant in this context is that GBIF has recently introduced a new data format intended for monitoring projects and sampling event protocols (GBIF 2019b).

© Pando F, Bonet F. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

In this presentation, we explore the suitability of GBIF data formats and workflows to serve LTER datasets, and the work it may take to transform typical LTER datasets into these formats. For this exercise, we use some datasets available via the DEIMS platform, corresponding to the same territory, (Sierra Nevada, Spain (e.g. Bonet 2016, Bonet 2018) and transform them into the GBIF's sample-based Event core publish them in the GBIF data network, and then perform an analysis to assess how the standardized datasets work in practice, both among themselves and also with typical “occurrence-based” GBIF datasets. Finally, we discuss our findings and make recommendations for the GBIF and LTER communities.

Keywords

Long-Term Ecological Research, LTER, environmental data, DEIMS-SDR, FAIR principles, GBIF

Presenting author

Francisco Pando

Presented at

Biodiversity_Next 2019

References

- Bonet F (2016) Meteorological Data of Sierra Nevada and surroundings. https://deims.org. URL: https://deims.org/activity/f0688bcc-9a91-4704-ac80-b5dda544dbd7
- Bonet F (2018) Soil moisture data from a Wireless Sensor Network of Sierra Nevada LTER platform. https://deims.org. URL: https://deims.org/dataset/d7f76ac5-8e30-4813-bdd7-6b1fd348a90e
- Fegraus E, Andelman S, Jones M, Schildhauer M (2005) Maximizing the Value of Ecological Data with Structured Metadata: An Introduction to Ecological Metadata Language (EML) and Principles for Metadata Creation. Bulletin of the Ecological Society of America 86 (3): 158-168. https://doi.org/10.1890/0012-9623(2005)86[158:mtvoed]2.0.co;2
- GBIF (2019a) The Global Biodiversity Information Facility. https://www.gbif.org/. Accessed on: 2019-4-03.
- GBIF (2019b) What is Darwin Core, and why does it matter? https://www.gbif.org/darwin-core. Accessed on: 2019-4-03.
- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLoS ONE 7 (1): e29715. https://doi.org/10.1371/journal.pone.0029715

- Wilkinson M, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J, da Silva Santos LB, Bourne P, Bouwman J, Brookes A, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo C, Finkers R, Gonzalez-Beltran A, Gray AG, Groth P, Goble C, Grethe J, Heringa J, 't Hoen PC, Hooft R, Kuhn T, Kok R, Kok J, Lusher S, Martone M, Mons A, Packer A, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S, Schultes E, Sengstag T, Slater T, Strawn G, Swertz M, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3: 160018. https://doi.org/10.1038/sdata.2016.18
- Wohner C, Peterseil J, Poursanidis D, Kliment T, Wilson M, Mirtl M, Chrysoulakis N (2019) DEIMS-SDR – A web portal to document research sites and their associated data. Ecological Informatics 51: 15-24. https://doi.org/10.1016/j.ecoinf.2019.01.005